

AD-A123 438

ASYMPTOTIC SPECTRAL ANALYSIS OF CROSS-PRODUCT MATRICES  
(U) PRINCETON UNIV NJ DEPT OF STATISTICS G S WATSON  
NOV 82 7 N00014-81-K-0146

1/1

UNCLASSIFIED

F/G 12/1

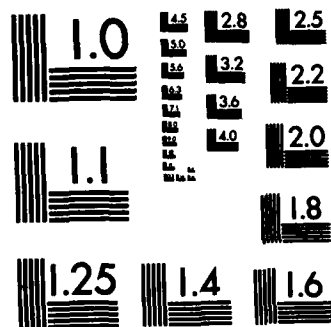
NL



END

FILMED

DTIC



MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

AD A 123 438

Report Number 7

## ASYMPTOTIC SPECTRAL ANALYSIS OF CROSS-PRODUCT MATRICES

G.S. Watson  
Department of Statistics  
Princeton University  
Princeton, N.J. 08544

NOVEMBER 1982

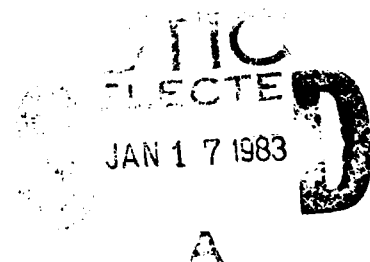
Prepared for

OFFICE OF NAVAL RESEARCH (Code 411SP)  
Statistics and Probability Branch  
Arlington, Virginia 22217  
under Contract N00014-81-K-0146  
SRO(103) Program in Non-Gaussian Signal Processing

S.C. Schwartz, Principal Investigator

Approved for public release; distribution unlimited

(12)



DTIC FILE COPY

88 01 17 049

| REPORT DOCUMENTATION PAGE   |                                     | READ INSTRUCTIONS<br>BEFORE COMPLETING FORM                                  |
|---|-------------------------------------|--|
| 1. REPORT NUMBER<br>7   | 2. GOVT ACCESSION NO.<br>AD A123438 | 3. RECIPIENT'S CATALOG NUMBER  |
| 4. TITLE (and Subtitle)<br>ASYMPTOTIC SPECTRAL ANALYSIS OF CROSS-<br>PRODUCT MATRICES   |                                     | 5. TYPE OF REPORT & PERIOD COVERED<br>Technical Report                       |
|   |                                     | 6. PERFORMING ORG. REPORT NUMBER   |
| 7. AUTHOR(s)<br>G.S. Watson   |                                     | 8. CONTRACT OR GRANT NUMBER(s)<br>N00014-81-K0146                            |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br>Department of Statistics<br>Princeton University<br>Princeton, NJ 08544  |                                     | 10. PROGRAM ELEMENT, PROJECT, TASK<br>AREA & WORK UNIT NUMBERS<br>NR SRO-103 |
| 11. CONTROLLING OFFICE NAME AND ADDRESS<br>Office of Naval Research (Code 411SP)<br>Department of the Navy<br>Arlington, Virginia 22217   |                                     | 12. REPORT DATE<br>November 1982   |
|   |                                     | 13. NUMBER OF PAGES<br>22  |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)   |                                     | 15. SECURITY CLASS. (of this report)<br>Unclassified                         |
|   |                                     | 15a. DECLASSIFICATION/DOWNGRADING<br>SCHEDULE                                |
| 16. DISTRIBUTION STATEMENT (of this Report)<br><br>Approved for public release; distribution unlimited  |                                     |  |
| 17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)  |                                     |  |
| 18. SUPPLEMENTARY NOTES   |                                     |  |
| 19. KEY WORDS (Continue on reverse side if necessary and identify by block number)<br><br>Spectral analysis<br>Asymptotic distribution<br>of covariance matrix  |                                     |  |
| 20. ABSTRACT (Continue on reverse side if necessary and identify by block number)<br><br>Let $x$ be a random vector in $IR^q$ and $M_n = n^{-1} \sum_{i=1}^n x_i x_i'$ be an estimator of $M = E x x'$ with spectral form $\sum \lambda_j P_j$ . An expository account is given of the estimation of $\lambda_j$ and $P_j$ from the eigenvalues and vectors of $M_n$ when $n$ is large. This includes a derivation of the basic formulae using a complex variable method in the book by<br>(over) |                                     |  |

(Abstract con't.)

Kato (1980) and a contrasting matrix method. Several extensions are indicated.

Accession For

NTIS GRA&I

DTIC TAB

Unannounced

Justification

DTIC

copy

inserted

2

A

# ASYMPTOTIC SPECTRAL ANALYSIS OF CROSS-PRODUCT MATRICES

by

G. S. Watson

Department of Statistics  
PRINCETON UNIVERSITY

## ABSTRACT

Let  $x$  be a random vector in  $\mathbb{R}^q$  and  $M_n = n^{-1} \sum_{i=1}^n x_i x_i'$  be an estimator of  $M = E x x'$  with spectral form  $\sum \lambda_j P_j$ . An expository account is given of the estimation of  $\lambda_j$  and  $P_j$  from the eigenvalues and and vectors of  $M_n$  when  $n$  is large. This includes a derivation of the basic formulae using a complex variable method in the book by Kato (1980) and a contrasting matrix method. Several extensions are indicated.

# ASYMPTOTIC SPECTRAL ANALYSIS OF CROSS-PRODUCT MATRICES

G. S. Watson

Princeton University

## 1. Introduction

T. W. Anderson (1963) derived the asymptotic distribution of the eigenvalues and vectors of the covariance matrix of a sample from a Gaussian distribution. Davis (1977) took his basic method and used it to get some results for the non-Gaussian case. The non-Gaussian case is of interest either because one wants to study the sensitivity of methods to deviations from Gaussianity - see e.g. Muirhead (1982) - or because one has to deal with other distributions. For example the distribution of the random vector might be entirely restricted to some manifold embedded in  $\mathbb{R}^q$  like the surface of the unit sphere or an hyperboloid of rotation; the cases of interest to us.

Kim (1978) at the suggestion of R. J. W. Beran, used results from the book by Kato (1976, 1980) on the perturbation theory of linear operators to find the asymptotic distribution of the eigenvalues of the matrix  $M_n = n^{-1} \sum_{i=1}^n x_i x_i'$  where the  $x_i$ 's are independently drawn from a certain distribution on the surface  $\Omega_q$  of the unit sphere in  $\mathbb{R}^q$ . Tyler (1979, 1981) also used Kato's method to get results in classical multivariate analysis. But the technique is not well-known, nor immediately evident from Kato's book.

Kato's method calls upon Cauchy's Theorem in Complex Variable Theory. Specifically consider the integral of  $(z - z_0)^p$  anti-clockwise around a simple closed curve  $C$  in the complex plane which does not go through  $z_0$ .

$$\int_C (z - z_0)^p dz$$

where  $p$  is an integer. Unless  $p = -1$ , it is always zero. When  $p = -1$ , it is zero if  $z_0$  is outside  $C$  and  $2\pi i$  where  $z_0$  is inside  $C$ .

The techniques and formulae below have many possible applications. Some are given in Watson (1982a) but most remain to be exploited. In the next section Kato's method is explained for symmetric non-random matrices and then applied in Section 3 to covariance matrices.

The key formulae in Section 2 and the results of Section 3 have of course been obtained before by direct matrix methods -

- though they are hard to justify. The Kato method not only gives a better insight but is easier to do and to extend, e.g. to provide asymptotic expansions.

## 2. The key to Kato

If  $T_0$  and  $T_1$  are real symmetric  $q \times q$  matrices and  $x$  is a small real number

$$T(x) = T_0 + xT_1 \tag{2.1}$$

can be thought of as a linear perturbation of  $T_0$ . Let the spectral



representation of the matrix  $T_0$  be

$$T_0 = \sum_{j=1}^r \lambda_j P_j, \quad r \leq q \quad (2.2)$$

where

$$\left. \begin{aligned} \lambda_1, \dots, \lambda_r \text{ are distinct real numbers,} \\ P_j' = P_j, \quad P_j P_k = \delta_{jk} P_j, \\ \text{rank } P_j = \text{trace } P_j = q_j, \quad \sum_{j=1}^r q_j = q \end{aligned} \right\} \quad (2.3)$$

Thus  $\lambda_j$  is an eigenvalue of  $T_0$  that is repeated  $q_j$  times. The invariant subspace  $V_j$  associated with  $\lambda_j$  has dimension  $q_j$  and  $P_j$  projects orthogonally onto  $V_j$  whose direct sum is  $\mathbb{R}^q$ .

The matrix  $T(x)$  may have  $q$  distinct eigenvalues but we would expect these to fall into  $r$  clusters about  $\lambda_1, \dots, \lambda_r$  and to condense on  $\lambda_1, \dots, \lambda_r$  as  $x \rightarrow 0$ . Equally the eigenvectors of  $T(x)$  should lead us to the eigen subspaces  $V_j$  as  $x \rightarrow 0$ . To show how this happens, define the resolvent of  $T_0$ ,  $R_0(\zeta)$  as

$$R_0(\zeta) = (T_0 - \zeta I_q)^{-1} \quad (2.4)$$

where  $\zeta$  is a complex number. By (2.2) we may write

$$R_0(\zeta) = \sum_{j=1}^r (\lambda_j - \zeta)^{-1} P_j \quad (2.5)$$

Observe that  $T_0$  and  $R_0(\zeta)$  commute.

If  $C$  is any contour in the complex plane which does not go through any  $\lambda_j$ , which are points on the real axis, Cauchy's Theorem and (2.5) imply that

$$\frac{1}{2\pi i} \int_C R_0(\zeta) d\zeta = \sum_{j \in C} P_j \quad (2.6)$$

where the sum is over the projectors  $P_j$  associated with eigenvalues  $\lambda_j$  inside  $C$ . The integral of a matrix is the matrix of integrals. Similarly

$$\frac{1}{2\pi i} \int_C T_0 R_0(z) dz = \sum_{j \in C} \lambda_j P_j \quad (2.7)$$

We observe that the trace of (2.6) gives the sum of the dimensions of the eigen subspaces associated with  $\lambda_j$  within  $C$ . Similarly the trace of (2.7) gives the sum of the eigen values (times their multiplicities) within  $C$ .

We now consider the resolvent of  $T(x)$ .

$$R(x, z) = (T(x) - zI_q)^{-1} = R_0(z)(I_q + xT_1R_0(z))^{-1} \quad (2.8)$$

If we apply the results of the previous paragraph to  $R(x, z)$ , we will get information about the eigen values  $\lambda(T(x))$  and projectors  $P(T(x))$ , of  $T(x)$ . As  $x \rightarrow 0$ , we would expect the values of  $\lambda(T(x))$  to condense on the eigen values  $\lambda_j$  of  $T_0$ .

To obtain the required formulae, we need to expand (2.8) as a power series. For a  $q \times q$  matrix  $A$ ,

$$(I_q + xA)^{-1} = I_q - xA + x^2A^2 - \dots \quad (2.9)$$

where the series is absolutely convergent provided  $|x| \|A\| < 1$ , where  $\|A\|$  is a norm of  $A$ . Thus we can say that for  $x$  sufficiently small,

$$(I_q + xA)^{-1} = I_q - xA + O(x^2). \quad (2.10)$$

Applying (2.10) to (2.8), we have, as  $|x| \rightarrow 0$ ,

$$R(x, z) = R_0(z) - x R_0(z) T_1 R_0(z) + O(x^2). \quad (2.11)$$

Consider now the analogue of (2.6) when  $C_j$  is a contour which encloses only the eigen value  $\lambda_j$ . Then

$$\frac{1}{2\pi i} \int_{C_j} R(x, \zeta) d\zeta = \frac{1}{2\pi i} \int_{C_j} R_0(\zeta) d\zeta + \frac{\pi}{2\pi i} \int_{C_j} R_0(\zeta) T_1 R_0(\zeta) d\zeta + O(x^2) \quad (2.12)$$

The first term on the right hand side (r.h.s.) of (2.12) is  $P_j$ .

To find the second term, we observe that, if we use (2.5) twice,

$$\begin{aligned} R_0(\zeta) T_1 R_0(\zeta) &= \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} \frac{P_k T_1 P_l}{(\lambda_k - \zeta)(\lambda_l - \zeta)} \\ &= \sum_{k=1}^{\infty} \frac{P_k T_1 P_k}{(\lambda_k - \zeta)^2} + \sum_{k < l} \frac{(P_k T_1 P_l + P_l T_1 P_k)}{\lambda_l - \lambda_k} \left( \frac{1}{\lambda_k - \zeta} - \frac{1}{\lambda_l - \zeta} \right) \end{aligned} \quad (2.13)$$

The contour integral of the first term on the r.h.s. of (2.13) is zero.

We get contributions from the second term when  $k$  or  $l$  equal  $j$  and they add to the symmetric matrix

$$\sum_{k \neq j} \frac{P_k T_1 P_j + P_j T_1 P_k}{\lambda_k - \lambda_j} \quad (2.14)$$

Observe, for later use, that this matrix has a zero trace because

$P_j P_k$  is null. Thus

$$\frac{1}{2\pi i} \int_{C_j} R(x, \zeta) d\zeta = P_j + x \sum_{k \neq j} \frac{P_k T_1 P_j + P_j T_1 P_k}{\lambda_j - \lambda_k} + O(x^2) \quad (2.15)$$

is the analogue of (2.6).

The analogue of (2.7) is obtained by integrating  $T(x)R(x, \zeta)$  which may, using (2.1) and (2.11), be written as

$$T(x)R(x, \zeta) = T_0 R_0(\zeta) + x(T_1 R_0(\zeta) - T_0 R_0(\zeta) T_1 R_0(\zeta)) + O(x^2). \quad (2.16)$$

The integral of the first term on the r.h.s. of (2.16) is that in (2.7).

To find the second and third terms we note that

$$T_1 R_0(z) = \sum_{k=1}^r (\lambda_k - z)^{-1} T_1 P_k,$$

$$T_0 R_0(z) T_1 R_0(z) = \sum_{k=1}^r \frac{\lambda_k P_k T_1 P_k}{(\lambda_k - z)^2} \quad (2.17)$$

$$+ \sum_{k < l} \frac{\lambda_k P_k T_1 P_l + \lambda_l P_l T_1 P_k}{(\lambda_l - \lambda_k)} \left( \frac{1}{\lambda_k - z} - \frac{1}{\lambda_l - z} \right), \quad (2.18)$$

where we have used (2.2) and (2.13). Thus we find that

$$\begin{aligned} \frac{1}{2\pi i} \int_{C_j} T(x) R(x, z) dz &= \lambda_j P_j + x(T_1 P_j + \sum_{k \neq j} \frac{\lambda_k P_k T_1 P_j + \lambda_j P_j T_1 P_k}{\lambda_j - \lambda_k}) \\ &+ O(x^2). \end{aligned} \quad (2.19)$$

Observe that the trace of the second term in the coefficient of  $x$  is zero. (2.19) is the analogue of (2.7).

In the applications we have in mind  $T(x)$  will have  $q$  distinct eigenvalues  $\lambda_1(x), \dots, \lambda_q(x)$  and (orthonormal) eigenvectors  $v_1(x), \dots, v_q(x)$  so that

$$T(x) = \sum_{i=1}^q \lambda_i(x) v_i(x) v_i(x)' \quad (2.20)$$

is the spectral form for  $T(x)$ . By using the reasoning that led to (2.6) and (2.7) we may then evaluate the l.h.s. of (2.15) and (2.19).

Thus

$$\frac{1}{2\pi i} \int_{C_j} R(x, z) dz = \sum v_i(x) v_i(x)' \quad (2.21)$$

and

$$\frac{1}{2\pi i} \int_{C_j} T(x) R(x, \zeta) d\zeta = \sum \lambda_i(x) v_i(x) v_i(x)' \quad (2.22)$$

where both sums are over  $i$  such that  $\lambda_i(x)$  are points inside the contour  $C_j$ .

Since  $\text{trace } v_i(x) v_i(x)' = v_i'(x) v_i(x) = 1$ , taking the trace of both sides of (2.15) and using (2.21) yields

$$\# \lambda_i(x) \text{ inside } C_j = q_j + O(x^2) \quad (2.23)$$

for any contour  $C_j$  enclosing  $\lambda_j$ . As  $x \rightarrow 0$ , one could use smaller and smaller contours. Hence as  $x \rightarrow 0$ , the eigenvalues of  $T(x)$  form clusters of  $q_j$  roots about  $\lambda_j$  ( $j = 1, \dots, r$ ) which condense upon  $\lambda_j$ . If we do not take the trace of (2.15) and write

$$\hat{P}_j = \sum_{\lambda_i(x) \in C_j} v_i(x) v_i(x)' \quad (2.24)$$

then (2.15) may be written as

$$\hat{P}_j = P_j + x \sum_{k \neq j} \frac{P_k T_j P_j + P_j T_j P_k}{\lambda_j - \lambda_k} + O(x^2) \quad (2.25)$$

Taking the trace of (2.19) yields

$$\sum \lambda_i(x) \text{ within } C_j = q_j \lambda_j + x \text{ trace } T_j P_j + O(x^2) \quad (2.26)$$

so dividing through by  $q_j$  and calling the l.h.s.  $\bar{\lambda}_j$ , the arithmetic mean of the  $j^{\text{th}}$  cluster, we have

$$\bar{\lambda}_j = \lambda_j + \frac{\pi}{q_j} \text{trace } T_1 P_j + O(x^2) \quad (2.27)$$

The formulae (2.25) and (2.27) are ideal for statistical applications, as will be seen in the next section. We close this section by observing that there is no problem except complexity in getting higher order approximations - one merely takes higher order terms in (2.10). For example, the coefficient of  $x^2$  in  $R(x_1)$  is  $R_0 T_1 R_0 T_1 R_0$  so using (2.5) and partial fraction expansions the contour integral may be evaluated to give a lengthy formula. One then finds that (2.23) may be improved to

$$\# \lambda_j(x) \text{ inside } C_j = q_j + O(x^3) \quad (2.28)$$

In Watson (1982a) explicit results are given when  $r = 2$ .

### 3. Large sample theory of symmetric cross-product matrices

Let  $x$  be a random vector in  $\mathbb{R}^q$  with components  $x^1, x^2, \dots, x^q$  and suppose that  $E x^i x^j x^k x^l$  exists for all  $i, j, k, l = 1, \dots, q$ . Let  $x'$  denote the transpose of  $x$ . Call  $E x x' = E[x^i x^j] = M$ , a symmetric  $q \times q$  matrix with spectral form

$$M = \sum_{j=1}^r \lambda_j P_j \quad (3.1)$$

If  $x_1, \dots, x_n$  are independent copies of  $x$ , define

$$M_n = n^{-1} \sum_{i=1}^n x_i x_i' \quad (3.2)$$

Then  $M_n \rightarrow M$  by the law of large numbers and by the multivariate central limit theorem

$$n^{1/2}(M_n - M) \xrightarrow{d} G. \quad (3.3)$$

The  $q(q+1)/2$  functionally independent elements of the symmetric matrix  $G$  are jointly Gaussian with zero means and a covariance matrix  $V$  whose elements are

$$E x^i x^j x^k x^l = E(x^i x^j) E(x^k x^l), \quad 1 \leq j, k \leq l. \quad (3.4)$$

To use the results of Section 2, we may write

$$M_n = M + n^{-1/2} \{n^{1/2}(M_n - M)\} \quad (3.5)$$

instead of

$$T(x) = T_0 + x T_1$$

From (3.3),  $T_1$  corresponds to  $G$ ,  $x$  to  $n^{-1/2}$ , and  $M$  to  $T_0$ . Provided no  $\lambda_j$  in (3.1) is zero, the matrix  $M_n$  will, with probability one, have distinct eigenvalues - Okamoto (1973). If say  $\lambda_1 = 0$ ,  $E(P_1 x)(P_1 x)'$  is a matrix of zeros so that, taking the trace,  $E \|P_1 x\|^2 = 0$ . Thus  $P_1 x$  is a null vector and  $M_n$  will have  $q_1$  zero roots and the data will determine the eigen subspace  $V_1$  exactly. This case has little interest so we assume that all the  $\lambda_j > 0$ .

The matrix  $M_n$  will be used to estimate the  $\lambda_j$  and  $P_j$ ,  $j=1, \dots, r$ . Combining (3.5) with (2.24), (2.25) and (2.27), we have the key results:

for  $j = 1, \dots, r$ ,

$$n^{1/2} (\hat{P}_j - P_j) \xrightarrow{d} \sum_{k \neq j} \frac{P_k G P_j + P_j G P_k}{\lambda_j - \lambda_k} \quad (3.6)$$

$$n^{1/2} (\bar{\lambda}_j - \lambda_j) \xrightarrow{d} \frac{1}{q_j} \text{trace } G P_j. \quad (3.7)$$

The r.h.s.'s of (3.6) and (3.7) are linear in the Gaussian matrix  $G$  so that the l.h.s.'s have asymptotically Gaussian distributions with zero means and variances and covariances that depend upon the covariance matrix  $V$  of  $G$ .

(3.7) is univariate and so easy to understand, e.g. it leads to a normal confidence interval for  $\lambda_j$ , although we will see that one will do better with a transformation. (3.6) describes the difference between estimated and true projectors and needs further simplification. Using the Euclidean matrix norm ( $\|A\|^2 = \text{trace } AA'$ ),

$$n \|\hat{P}_j - P_j\|^2 \xrightarrow{d} 2 \sum_{k \neq j} \frac{\text{trace } P_j G P_k G}{(\lambda_j - \lambda_k)^2} \quad (3.8)$$

Again, one might examine the different effects of  $\hat{P}_j$  and  $P_j$  on vectors. For example, if  $v \in V_j$ ,

$$n^{1/2}(\hat{P}_j v - P_j v) \xrightarrow{d} \sum_{k \neq j} \frac{P_k G v}{\lambda_j - \lambda_k} \quad (3.9)$$

so

$$n \|\hat{P}_j v - P_j v\|^2 \xrightarrow{d} \sum_{k \neq j} \frac{v' G P_k G v}{\lambda_j - \lambda_k} \quad (3.10)$$

More fundamentally if  $\hat{V}_j$  is the subspace onto which  $\hat{P}_j$  projects,  $\hat{V}_j$  will be "close" to  $V_j$  if  $\cos \theta = v' \hat{v}$  is always large when  $v \in V_j$  and  $\hat{v} \in \hat{V}_j$ ,  $\|v\| = 1$ ,  $\|\hat{v}\| = 1$ . Thus we should seek the stationary values of  $(P_j u)'(\hat{P}_j w)$ , given  $\|P_j u\| = \|\hat{P}_j w\| = 1$ , i.e., we should consider

$$2u'P_j\hat{P}_jw - \theta u'P_ju - \phi w'\hat{P}_jw$$



where  $\theta$  and  $\phi$  are Lagrangian multipliers. Hence

$$\left. \begin{aligned} P_j \hat{P}_j w - \theta P_j u &= 0, \\ \hat{P}_j P_j u - \phi \hat{P}_j w &= 0, \end{aligned} \right\} \quad (3.11)$$

so that

$$\begin{aligned} \theta = \phi &= \text{stationary value of } (P_j u)' (\hat{P}_j w) \\ &= C, \text{ say.} \end{aligned}$$

Hence the equations (3.11) will only have a solution if

$$\begin{vmatrix} -C P_j & P_j \hat{P}_j \\ \hat{P}_j P_j & -C \hat{P}_j \end{vmatrix} = 0 \quad (3.12)$$

This equation for  $C$  may be reduced to

$$\left| P_j \hat{P}_j P_j - C^2 P_j \right| = 0 \quad (3.13)$$

which has  $q_j$  non-zero roots  $C_{j1}^2$ . If, however, (3.6) is used, one finds eventually that all the  $C_{j1}^2$  are unity. Watson (1982a) deals with the case where  $r = 2$  and shows, by taking the next term in the expansions in Section 2, that  $n(1 - C_{j1}^2)$  have asymptotic distributions. It is conjectured that for any  $r$  the asymptotic joint distribution of  $n(1 - C_{j1}^2)$ , ...,  $n(1 - C_{qj1}^2)$  is the joint distribution of the non-zero eigenvalues of

$$\sum_{k \neq j} \frac{P_{j1} G P_k G P_{j1}}{(\lambda_j - \lambda_k)^2} \quad (3.14)$$

Some of the above results become easier to understand if we write, since  $I_q = P_1 + \dots + P_r$ ,

$$y_j = P_j x, \quad x = y_1 + \dots + y_r \quad (3.15)$$

One of the reasons results become simpler for the Gaussian is that there  $y_1, \dots, y_r$  are independent. Since  $Exx' = M = \lambda_j P_j$ ,

$$\left. \begin{aligned} Ey_k y_l' &= 0, \quad l \neq k, \quad Ey_j y_j' = \lambda_j P_j \\ Ey_l' y_k &= 0, \quad (l \neq k), \quad Ey_j' y_j = \lambda_j q_j \end{aligned} \right\} \quad (3.16)$$

Thus (3.7) may be rewritten as

$$\begin{aligned} n^{1/2} (\bar{\lambda}_j - \lambda_j) &\sim \frac{n^{1/2}}{q_j} \text{trace} \left( \frac{1}{n} \sum_{i=1}^n y_{ji} y_{ji}' - \lambda_j P_j \right) \\ &= \frac{n^{1/2}}{q_j} \left( \frac{1}{n} \sum_{i=1}^n y_{ji}' y_{ji} - \lambda_j q_j \right) \end{aligned} \quad (3.17)$$

so that by (3.16) and the Central Limit Theorem

$$\mathcal{L} \quad n^{1/2} (\bar{\lambda}_j - \lambda_j) \longrightarrow G_1(0, \text{var}(y_j' y_j) q_j^{-2}) \quad (3.18)$$

where  $G_q(\mu, \Sigma)$  stands for the Gaussian distribution in  $q$  dimensions with mean vector  $\mu$  and covariance matrix  $\Sigma$ . Similarly (3.6) can be written as

$$n^{1/2} (\hat{P}_j - P_j) \sim n^{-1/2} \sum_{i=1}^n \sum_{k \neq j} \frac{y_{ki} y_{ji}' + y_{ji} y_{ki}'}{\lambda_j - \lambda_k} \quad (3.19)$$

If  $\mathcal{L} x = G_q(0, M)$ ,  $y_1, y_2, \dots, y_r$  are independent and  $\mathcal{L} y_j' y_j \lambda_j^{-1} = \chi_{q_j}^2$  so that  $\text{var}(y_j' y_j) = \lambda_j^2 2q_j$ . Then (3.18) reads:

$$\mathcal{L} n^{1/2} (\bar{\lambda}_j - \lambda_j) \longrightarrow G_1(0, 2\lambda_j^2/q_j)$$

Hence

$$\mathcal{L} n^{1/2} (\log(\bar{\lambda}_j/\lambda_j) - 1) \longrightarrow G_1(0, 2/q_j) \quad (3.20)$$

giving the variance stabilizing transformation. Moreover in this Gaussian case the  $n^{1/2} (\bar{\lambda}_j - \lambda_j)$  or  $n^{1/2} (\log(\bar{\lambda}_j/\lambda_j) - 1)$  are asymptotically independent, a simplifying result which is not true in general. Under no circumstances could it be expected that the  $\hat{P}_j$  would be independent since  $\sum_{j=1}^r \hat{P}_j = I_q$ .

With this introduction, the compact paper by Tyler (1981) may be read for more details on  $\hat{P}_j$ . He also gives tests. special case of  $r = 2$  and distributions restricted to  $\Omega_q$ , Watson (1982a).

If an additional assumption is made that the distribution of  $x$  depends only upon  $\|y_1\|, \dots, \|y_r\|$  more results may be derived - see Watson 1982b).

#### 4. Direct approach to large sample theory of cross-product matrices.

The eigenvalues of  $M_n$  are the roots  $\hat{\lambda}$  of

$$|M_n - \lambda I| = |M - \lambda I + \frac{1}{\sqrt{n}} T_1| = 0 \quad (4.1)$$

where, as in Section 3,  $M = \sum_{j=1}^r \lambda_j P_j$ ,  $T_1 = \sqrt{n}(M_n - M)$ . Suppose orthonormal eigenvectors are selected to span each of the invariant

subspaces  $V_j$  and arranged as column vectors to form a  $q \times q$  orthogonal matrix  $H$ . Let the first  $q_1$  columns correspond to  $V_1$ , the next  $q_2$  columns to  $V_2$ , etc., and write it in partitioned form

$$H = [H_1, \dots, H_r] \quad (4.2)$$

Then since  $H'H = HH' = I_q$ , we have

$$\begin{aligned} H_a'H_b &= O(a \neq b), \quad H_a'H_a = I_{q_a}, \\ H_1H_1' + \dots + H_rH_r' &= I_q, \end{aligned} \quad (4.3)$$

$$H_aH_a' = P_a, \quad a = 1, \dots, r,$$

and

$$H'MH = D(\lambda_j I_{q_j}), \quad (4.4)$$

a matrix partitioned so all  $r^2$  submatrices are zero except for the multiples of identity matrices on the diagonal.

If  $H$  is applied to (4.1), it takes the partitioned form.

$$\left| (\lambda_1 - \lambda) I_{q_1} \delta_{1j} + n^{-1/2} H_1'T_1H_j \right| = 0 \quad (4.5)$$

Since  $n \rightarrow \infty$ , we seek the  $O(1)$  and  $O(n^{-1/2})$  terms only in the expansion of (4.5). Applying the formula

$$\begin{vmatrix} A & C \\ B & D \end{vmatrix} = |A| |D - BA^{-1}C|$$

when  $A$  is the leading submatrix of (4.5), it is seen that  $BA^{-1}C$  is  $O(n^{-1})$  and so negligible. Hence we may repeat the procedure to find that equation (4.5) and hence (4.1) is, to this order

$$\prod_{j=1}^r |(\lambda_j - \lambda) I_{q_j} + n^{-1/2} H_j' T_1 H_j| = 0 \quad (4.6)$$

This shows that the eigenvalues of  $M_n$ , for large  $n$ , form clusters about the  $r$  distinct roots  $\lambda_j$  of  $M$ . Expanding the  $j^{\text{th}}$  factor in (4.6) to  $O(n^{-1/2})$ , we only need the product of the diagonal terms and find the equation

$$(\lambda_j - \lambda)^{q_j} (1 + n^{-1/2} \frac{\text{trace } H_j' T_1 H_j}{\lambda_j - \lambda}) = 0 \quad (4.7)$$

Since  $\text{trace } H_j' T_1 H_j = \text{trace } H_j H_j' T_1 = \text{trace } P_j T_1$  by (4.3) the  $q_j$  roots of (4.7) tend to  $\lambda_j$  as  $n \rightarrow \infty$  and the leading terms of the polynomial (degree  $q_j$ ) equation for  $\lambda$  are

$$\lambda^q - \{q\lambda_j + n^{-1/2} \text{trace } P_j T_1\} \lambda^{q-1} + \dots = 0 \quad (4.8)$$

so that if the roots of this equation are denoted by  $\hat{\lambda}$ ,

$$\Sigma \hat{\lambda} = q\lambda_j + n^{-1/2} \text{trace } P_j T_1, \quad (4.9)$$

as we found in (3.27). But one cannot expect to obtain the roots in the cluster for  $\lambda_j$  from (4.7) (it gives them to be  $\lambda_j$  ( $q-1$  times),  $\lambda_j + n^{-1/2} \text{trace } P_j T_1$  (once)) because when  $\lambda$  is within  $n^{-1/2}$  of  $\lambda_j$  all the terms in the matrix in (4.6) are of order  $n^{-1/2}$ . However, (4.8) does give the correct coefficient for  $\lambda^{q-1}$  in (4.6).

Since the eigenvalues of  $M_n$  will in general be distinct, the approximations made above are inadequate to discuss e.g. the joint distribution of the eigenvalues in a cluster.

The following direct derivation of the analogue of (2.25) or (3.6) is harder to justify. Write

$$\hat{P}_j = P_j + n^{-1/2}\Delta, \Delta = n^{1/2}(\hat{P}_j - P_j) \quad (4.10)$$

and, because the roots in the cluster are within  $n^{-1}$  of  $\lambda_j$ , set

$$M_n \hat{P}_j = (\lambda_j + n^{-1/2}\delta)\hat{P}_j \quad (4.11)$$

i.e.

$$(M + n^{-1/2}T_1)(P_j + n^{-1/2}\Delta) = (\lambda_j + n^{-1/2}\delta)(P_j + n^{-1/2}\Delta)$$

so that the terms in  $n^{-1/2}$  yield the equation

$$T_1 P_j + M\Delta = \lambda_j \Delta + \delta P_j$$

or

$$(M - \lambda_j I)\Delta = -T_1 P_j + \delta P_j \quad (4.12)$$

But  $M - \lambda_j I = \sum_{k \neq j} (\lambda_k - \lambda_j) P_k$  so that we could replace  $\Delta$  by  $\Delta - P_j X$  and still satisfy (4.12). Thus (4.12) is solved by multiplying (4.12) by  $\sum_{k \neq j} (\lambda_k - \lambda_j)^{-1} P_k$  and adding  $P_j X$  so that

$$\Delta = \sum_{k \neq j} (\lambda_j - \lambda_k)^{-1} P_k T_1 P_j + P_j X.$$

However, from (4.10) we see that  $\Delta$  must be symmetric and this determines  $X$ . Thus

$$\Delta = n^{1/2}(\hat{P}_j - P_j) = \sum_{k \neq j} \frac{P_k T_1 P_j + P_j T_1 P_k}{\lambda_j - \lambda_k} \quad (4.13)$$

which is the desired result. However, without the results of Section 3, (4.10) and (4.11) are merely intuitions.

### 5. Additional Remarks

(i) Suppose  $(x_i, y_i)$  for  $i = 1, \dots, n$  are independent copies of a pair of random vectors  $(x, y)$ ,  $x \in \mathbb{R}^q$ ,  $y \in \mathbb{R}^p$ . Define the estimator of  $N = Exy'$

$$N_n = n^{-1} \sum_{i=1}^n x_i y_i' \quad (5.1)$$

and assume that the Central Limit Theorem holds so that

$$n^{1/2} (N_n - N) \xrightarrow{d} F \quad (5.2)$$

One may wish to estimate the singular values of  $N$  so one would find the non-zero eigenvalues of  $N_n N_n'$  or  $N_n' N_n$  whichever is the smaller. But using (5.2),

$$N_n' N_n \sim N' N + n^{-1/2} (F' N + N' F) \quad (5.3)$$

so that the previous theory is applicable.

(ii) Suppose that  $x_1, \dots, x_n$  is a sample from one  $q$  dimensional distribution,  $x_1^*, \dots, x_m^*$  a sample from another distribution in  $\mathbb{R}^q$  and let  $M_n = n^{-1} \sum x_i x_i'$ ,  $M_m^* = n^{-1} \sum x_i^* x_i^{*'}.$  Then we often need to study the solutions of

$$(M_n - \lambda M_m^*)v = 0 \quad (5.4)$$

Anderson (1958) gives examples and calls  $\lambda$  and  $v$  the eigen values and vectors of  $M_n$  in the metric of  $M_n$ . To study them in large samples we suppose that the Central Limit Theorem applies in both cases so that as  $m, n \rightarrow \infty$

$$\left. \begin{aligned} M_n &\sim M + n^{-1/2} G, \\ M_n^* &\sim M^* + n^{-1/2} G^* \end{aligned} \right\} \quad (5.5)$$

The eigenvalues are those of  $M_n^{*-1/2} M_n M_n^{*-1/2}$  which by (5.5) are those of

$$\begin{aligned} &M^{*-1/2} \left( I_q - \frac{n}{2} G^* M^{*-1} \right) (M + n^{-1/2} G) \left( I_q - \frac{n}{2} M^{*-1} G^* \right) M^{*-1/2} \\ &= M^{*-1/2} M M^{*-1/2} + M^{*-1/2} G M^{*-1/2} \\ &\quad - \frac{n}{2} M^{*-1/2} G^* M^{*-1} M M^{*-1/2} \\ &\quad - \frac{n}{2} M^{*-1/2} M M^{*-1} G^* M^{*-1/2} \end{aligned} \quad (5.6)$$

If we set  $n = \alpha l$ ,  $m = \beta l$  with  $\alpha, \beta > 0$  and  $l \rightarrow \infty$  (5.6) has the form of a symmetric fixed matrix plus  $l^{-1/2}$  times a symmetric Gaussian matrix so that the earlier theory is applicable.



References

- Anderson, T.W. (1958), Introduction to Multivariate Analysis,  
John Wiley & Sons, New York.
- Anderson, T.W. (1963), Asymptotic Theory for Principal Components,  
Ann.Math.Stat. 34, 122-148.
- Davis, A.W. (1977), Asymptotic Theory for Principal Components,  
Austral.J.Statist. 19 (3), 206-212.
- Kim, K.M. (1978), Orientation Shift Model on the Sphere.  
Ph.D. Thesis, Univ. of California, Berkeley.
- Kato, T. (1980), Perturbation Theory for Linear Operators,  
2nd Edition, Springer-Verlag, New York.
- Muirhead, Robb J. (1982), Aspects of Multivariate Statistical Theory,  
John Wiley & Sons, New York.
- Tyler, D.E. (1979), Redundancy Analysis and Associated Asymptotic  
Distribution Theory, Ph.D. Thesis, Princeton  
University.
- Tyler, D.E. (1981), Asymptotic Inference for Eigenvectors,  
Ann.Statistics 9 (4), 725-736.
- Watson, G.S. (1982a), Large Sample Theory for Distributions on the  
Hypersphere with Rotational Symmetries.  
Technical Report, Dept. of Statistics,  
Princeton University.
- Watson, G.S. (1982b), Distributions in  $\mathbb{R}^q$  with Rotational Symmetries.  
Technical Report, Dept. of Statistics,  
Princeton University.

OFFICE OF NAVAL RESEARCH  
STATISTICS AND PROBABILITY PROGRAM

BASIC DISTRIBUTION LIST  
FOR  
UNCLASSIFIED TECHNICAL REPORTS

FEBRUARY 1982

| Copies  | Copies  |
|---|---|
| Statistics and Probability<br>Program (Code 411(SP))<br>Office of Naval Research<br>Arlington, VA 22217 3   | Navy Library<br>National Space Technology Laboratory<br>Attn: Navy Librarian<br>Bay St. Louis, MS 39522 1                               |
| Defense Technical Information<br>Center<br>Cameron Station<br>Alexandria, VA 22314 12   | U. S. Army Research Office<br>P.O. Box 12211<br>Attn: Dr. J. Chandra<br>Research Triangle Park, NC<br>27706 1                           |
| Commanding Officer<br>Office of Naval Research<br>Eastern/Central Regional Office<br>Attn: Director for Science<br>Barnes Building<br>495 Summer Street<br>Boston, MA 02210 1 | Director<br>National Security Agency<br>Attn: R51, Dr. Maar<br>Fort Meade, MD 20755 1   |
| Commanding Officer<br>Office of Naval Research<br>Western Regional Office<br>Attn: Dr. Richard Lau<br>1030 East Green Street<br>Pasadena, CA 91101 1                          | ATAA-SL, Library<br>U.S. Army TRADOC Systems<br>Analysis Activity<br>Department of the Army<br>White Sands Missile Range, NM<br>88002 1 |
| U. S. ONR Liaison Office - Far East<br>Attn: Scientific Director<br>APO San Francisco 96503 1   | ARI Field Unit-USAREUR<br>Attn: Library<br>c/o ODCSPER<br>HQ USAEREUR & 7th Army<br>APO New York 09403 1                                |
| Applied Mathematics Laboratory<br>David Taylor Naval Ship Research<br>and Development Center<br>Attn: Mr. G. H. Gleissner<br>Bethesda, Maryland 20084 1                       | Library, Code 1424<br>Naval Postgraduate School<br>Monterey, CA 93940 1   |
| Commandant of the Marine Corps<br>(Code AX)<br>Attn: Dr. A. L. Slafkosky<br>Scientific Advisor<br>Washington, DC 20380 1  | Technical Information Division<br>Naval Research Laboratory<br>Washington, DC 20375 1   |
|   | OASD (I&L), Pentagon<br>Attn: Mr. Charles S. Smith<br>Washington, DC 20301 1  |

## Copies

Director  
AMSAA  
Attn: DRXS-MP, H. Cohen  
Aberdeen Proving Ground, MD 1  
21005

Dr. Gerhard Heiche  
Naval Air Systems Command  
(NAIR 03)  
Jefferson Plaza No. 1  
Arlington, VA 20360 1

Dr. Barbara Bailar  
Associate Director, Statistical  
Standards  
Bureau of Census  
Washington, DC 20233 1

Leon Slavin  
Naval Sea Systems Command  
(NSEA 05H)  
Crystal Mall #4, Rm. 129  
Washington, DC 20036 1

B. E. Clark  
RR #2, Box 647-B  
Graham, NC 27253 1

Naval Underwater Systems Center  
Attn: Dr. Derrill J. Bordelon  
Code 601  
Newport, Rhode Island 02840 1

Naval Coastal Systems Center  
Code 741  
Attn: Mr. C. M. Bennett  
Panama City, FL 32401 1

Naval Electronic Systems Command  
(NELEX 612)  
Attn: John Schuster  
National Center No. 1  
Arlington, VA 20360 1

Defense Logistics Studies  
Information Exchange  
Army Logistics Management Center  
Attn: Mr. J. Dowling  
Fort Lee, VA 23801 1

## Copies

Reliability Analysis Center (RAC)  
RADC/RBRAC  
Attn: I. L. Krulac  
Data Coordinator/  
Government Programs  
Griffiss AFB, New York 13441 1

Technical Library  
Naval Ordnance Station  
Indian Head, MD 20640 1

Library  
Naval Ocean Systems Center  
San Diego, CA 92152 1

Technical Library  
Bureau of Naval Personnel  
Department of the Navy  
Washington, DC 20370 1

Mr. Dan Leonard  
Code 8105  
Naval Ocean Systems Center  
San Diego, CA 92152 1

Dr. Alan F. Petty  
Code 7930  
Naval Research Laboratory  
Washington, DC 20375 1

Dr. M. J. Fischer  
Defense Communications Agency  
Defense Communications Engineering  
Center  
1860 Wiehle Avenue  
Reston, VA 22090 1

Mr. Jim Gates  
Code 9211  
Fleet Material Support Office  
U. S. Navy Supply Center  
Mechanicsburg, PA 17055 1

Mr. Ted Tupper  
Code M-311C  
Military Sealift Command  
Department of the Navy  
Washington, DC 20390 1

Copies

Copies

Mr. F. R. Del Priori  
Code 224  
Operational Test and Evaluation  
Force (OPTEVFOR)  
Norfolk, VA 23511

1

END

FILMED

2-83

DTIC